

GENERALIZED LAMINAR FAMILIES AND CERTAIN FORBIDDEN MATRICES

Peter J. Dukes

June 13, 2014

ABSTRACT. Recall that in a laminar family, any two sets are either disjoint or contained one in the other. Here, a parametrized weakening of this condition is introduced. Let us say that a set system $\mathcal{F} \subseteq 2^X$ is t -laminar if $A, B \in \mathcal{F}$ with $|A \cap B| \geq t$ implies $A \subseteq B$ or $B \subseteq A$. We obtain very close asymptotic bounds in terms of n on the maximum size of a 2-laminar family $\mathcal{F} \subseteq 2^{[n]}$. A construction for 3-laminar families and a crude analysis for general t are also given.

1. Introduction

Let X be a set. A *laminar* family of subsets $\mathcal{L} \subseteq 2^X$ has the property that for $A, B \in \mathcal{L}$, either $A \subseteq B$, $B \subseteq A$, or $A \cap B = \emptyset$. Consider a finite set X , say of cardinality n . (Usually $X = [n] := \{1, \dots, n\}$ is assumed.) Then $|\mathcal{L}| \leq 2n$ by an easy (strong) induction argument. Equality is achieved if \mathcal{L} consists of a full chain of size $n + 1$, along with singletons.

Said another way, \mathcal{L} is laminar if $A, B \in \mathcal{L}$ with $|A \cap B| \geq 1$, implies $A \subseteq B$ or $B \subseteq A$. Now, let t be a positive integer and say a family $\mathcal{F} \subseteq 2^X$ is t -laminar if, whenever $A, B \in \mathcal{F}$ with $|A \cap B| \geq t$, we have $A \subseteq B$ or $B \subseteq A$. It is clear that an s -laminar family is also t -laminar for $t > s$.

From the point of view of obtaining upper bounds on t -laminar families \mathcal{F} , there is no loss in assuming $\binom{[n]}{\leq t} \subseteq \mathcal{F}$. Also, the set $[n]$ itself can be thrown into \mathcal{F} for free.

The author's original motivation for this topic was not to generalize laminar families. It originated from a certain problem of Richard Anstee on forbidden configurations in zero-one matrices, [1].

In this direction, say that a zero-one matrix M *avoids* another zero-one matrix Z if no permutation of rows and columns of Z appears as a submatrix of M . This is especially interesting when M has a fixed number n of columns and both M and Z have no repeated rows. (It should be clarified that Anstee works with the transpose: m rows and no repeated columns, but the difference is not too important here.) Under the assumption of distinct rows, it follows that sufficiently large M are guaranteed to contain a copy of Z . The key question is how many rows M can have while avoiding Z .

The rows of M can be interpreted as characteristic vectors of distinct subsets of $[n]$. Likewise, a set system $\mathcal{F} \subseteq 2^{[n]}$ defines a $|\mathcal{F}| \times n$ matrix $M(\mathcal{F})$ with rows indexed by \mathcal{F} and

$$M(\mathcal{F})_{A,i} = \begin{cases} 1 & \text{if } i \in A \in \mathcal{F}, \\ 0 & \text{if } i \notin A. \end{cases}$$

Research of the author is supported by NSERC.

Now consider the matrix

$$Z = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}.$$

We have chosen $Z = F_{0112}^\top$ from [1], where in that more general context the subscripts denote the number of rows (our columns) which are 00, 01, 10, and 11. Incidentally, among all 2×4 matrices, the forbidden configuration problem is by far least understood for our Z .

Observe that $M(\mathcal{F})$ contains Z as a (possibly permuted) submatrix if and only if there exist four points $w, x, y, z \in [n]$ and two sets $A, B \in \mathcal{F}$ so that

- $x, y, z \in A$ but $w \notin A$; and
- $w, y, z \in B$ but $x \notin B$.

It follows that forbidding Z corresponds to building a 2-laminar family. More generally, forbidding the $2 \times (t+2)$ matrix F_{011t}^\top corresponds with t -laminar families for each positive integer t .

In another direction, let us recall that $2^{[n]}$ is a graded poset. In a laminar family, we must have that every rank one element (singleton set) belongs to a unique maximal chain in the induced ordering. Similarly, in a t -laminar family, every element of rank t sits below a unique maximal chain. This is potentially a natural property for other graded posets; e.g. in the number-theoretic setting ($[n]$ with divisibility) or finite geometric setting (\mathbb{F}_q^k with subspace containment).

Summarizing the above, we claim a collection of equivalent properties for set systems.

Proposition 1.1. *The following are equivalent for $\mathcal{F} \subseteq 2^{[n]}$:*

- (1) \mathcal{F} is t -laminar;
- (2) the matrix $M(\mathcal{F})$ avoids F_{011t}^\top ;
- (3) every element of $\mathcal{F} \cap \binom{[n]}{t}$ belongs to a unique maximal chain in $\mathcal{F} \cap \binom{[n]}{\geq t}$.

In this note, we are especially interested in asymptotic upper bounds on such families \mathcal{F} for the case $t = 2$ and $n \gg 0$. We obtain fairly close estimates as our main result.

Theorem 1.2. *Let $f(n) + n + 1$ be the maximum size of a 2-laminar family on an n -set. Then*

$$\begin{aligned} \text{(a)} \quad \liminf_{n \rightarrow \infty} f(n) \binom{n}{2}^{-1} &\geq 1.3818, \text{ and} \\ \text{(b)} \quad \limsup_{n \rightarrow \infty} f(n) \binom{n}{2}^{-1} &\leq 1.3821. \end{aligned}$$

As the reader might expect, the lower bound follows from a constructive argument. The upper bound is obtained from a recursive linear program. These define Sections 2 and 3 to follow. Interestingly, obtaining even slightly better bounds is probably very difficult; indeed, determining an exact limit as above is essentially as hard as settling certain open cases for projective planes. We are also able to say something about the case $t = 3$, but (for now) this represents a barrier.

2. The lower bound: designs and packings

A t -wise balanced design is a pair (V, \mathcal{B}) , where V is a set of v points and \mathcal{B} is a collection of subsets of V called blocks (and having size $\geq t$) such that every t -subset of points belongs to exactly one

block. A t -wise packing is defined similarly, except that some t -subsets of points may belong to zero blocks.

If the block sizes are a constant k , we speak of a t -($v, k, 1$) design (or packing). Here, the ‘1’ denotes the (maximum) number of blocks containing a fixed t -subset; in more general settings this parameter may be any nonnegative integer λ .

The connection with t -laminar families is straight from definitions.

Proposition 2.1. *Let $\mathcal{F} \subseteq 2^{[n]}$ be a t -laminar family. Then any antichain in \mathcal{F} forms the blocks of a t -wise packing on n points. In particular, the maximal sets in $\mathcal{F} \setminus \{[n]\}$ form a t -wise packing and the sets in \mathcal{F} of fixed cardinality $k > t$ in \mathcal{F} form a t -($n, k, 1$) packing.*

A t -laminar family can arise from ‘nested’ packings.

Construction 2.2. *Suppose $([n], \mathcal{B})$ is a t -wise packing. Define \mathcal{F} by replacing each block $K \in \mathcal{B}$ by any copy of a t -laminar family on K . Then \mathcal{F} is t -laminar with ground set $[n]$.*

PROOF. For $K \in \mathcal{B}$, let \mathcal{E}_K denote the t -laminar family replacing K in the construction, so that $\mathcal{F} = \cup_{K \in \mathcal{B}} \mathcal{E}_K$. Consider $A, B \in \mathcal{F}$. If $A, B \in \mathcal{E}_K$ for some K , it follows that $A \subseteq B$, $B \subseteq A$, or $|A \cap B| < t$. On the other hand, if A and B each belong to different \mathcal{E}_{K_1} and \mathcal{E}_{K_2} , then $|K_1 \cap K_2| < t$ since \mathcal{B} is a t -packing. It follows that $|A \cap B| < t$ in this case. \square

In the notation of our main result, Theorem 1.2, recall that we are letting $f(n)$ denote the maximum $|\mathcal{F} \cap \binom{[n]}{\geq 2}|$ over all 2-laminar families \mathcal{F} .

Corollary 2.3. *Suppose there exists a 2-($n, m, 1$) packing with b blocks. Then $f(n) \geq b \cdot f(m) + 1$.*

Note the ‘+1’ is due to the universal set $[n]$. Now, Construction 2.2 can be used inductively. The choice of packings (or designs) and block sizes must now be explored further. To this end, we recall two famous families of designs.

Lemma 2.4. *There exists a 2-($q^2, q, 1$) design and a 2-($q^2 + q + 1, q + 1, 1$) design for every prime power q .*

The designs stipulated in Lemma 2.4 are, respectively, the affine and projective planes of order q . In particular, the projective plane for $q = 2$ is the well known *Fano plane*, a 2-(7, 3, 1)-design. They lead to an easy construction of 2-laminar families for an infinite sequence of n .

Proposition 2.5. *Let $n = 7^{2^r}$. Then*

$$(2.1) \quad f(n) \geq \binom{n}{2} \left[1 + \frac{1}{\binom{3}{2}} + \frac{1}{\binom{7}{2}} + \frac{1}{\binom{49}{2}} + \cdots + \frac{1}{\binom{n}{2}} \right].$$

PROOF. Use induction on r . If $r = 0$, then $n = 7$ and we construct a family $\mathcal{F}_0 := \binom{[7]}{2} \cup \mathcal{B} \cup \{[7]\}$, where $\mathcal{B} \subseteq \binom{[7]}{3}$ are the $\binom{7}{2}/\binom{3}{2}$ blocks of a Fano plane.

It is easy to check that \mathcal{F}_0 is 2-laminar, leading to $f(7) \geq 21 + 7 + 1 = 29$. Now for $i \geq 1$, construct \mathcal{F}_i from \mathcal{F}_{i-1} by applying Construction 2.2 to the 2-($7^{2^i}, 7^{2^{i-1}}, 1$) design in Lemma 2.4. It follows that \mathcal{F}_r is 2-laminar with ground set $[n]$. It remains to count $|\mathcal{F}_r|$.

In general, there are $\binom{n}{2}\binom{m}{2}^{-1}$ blocks in a $2-(n, m, 1)$ design. With $m = 7^{2^{r-1}}$ and $n = m^2$, we have

$$\frac{f(n)}{\binom{n}{2}} \geq \frac{|\mathcal{F}_r|}{\binom{n}{2}} = \frac{|\mathcal{F}_{r-1}|}{\binom{m}{2}} + \frac{1}{\binom{n}{2}}.$$

We obtain the (2.1) after an easy induction. \square

To obtain a construction for general $n \gg 0$, we make use of an asymptotic result on maximum 2-packings.

Lemma 2.6 ([2]). *Given any k , there exists a universal constant c_k such that, for all n , the maximum size of a $2-(n, k, 1)$ packing is at least $\lfloor \frac{n}{k} \lfloor \frac{n-1}{k-1} \rfloor \rfloor - c_k$.*

We are now ready to prove the first half of our main result.

PROOF OF THEOREM 1.2(a). Let $k = 7^4$. It follows from Proposition 2.5 and a quick calculation that $f(k) \geq 1.3818\binom{k}{2}$. Now, take an $(n, k, 1)$ -packing, which by Lemma 2.6 has at least $b = \binom{n}{2}\binom{k}{2}^{-1} - o(n^2)$ blocks. It follows from Corollary 2.3 that $f(n) > [1.3818 - o(1)]\binom{n}{2}$, as desired. \square

For $t = 3$, there is a classical family of designs we can use to take the place of the finite planes.

Lemma 2.7. *There exists a $3-(q^2 + 1, q + 1, 1)$ design for every prime power q .*

These designs are known as *circle geometries*. They arise from the action of $\text{PGL}(2, q^2)$ on $\mathbb{F}_q \cup \{\infty\}$ with its natural containment in $\mathbb{F}_{q^2} \cup \{\infty\}$. Using the same construction as for $t = 2$, we have the following.

Proposition 2.8. *Let $n = 3^{2^r} + 1$. There exists a 3-laminar family of $[n]$ with size*

$$(2.2) \quad 1 + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} \left[1 + \frac{1}{\binom{4}{3}} + \frac{1}{\binom{10}{3}} + \cdots + \frac{1}{\binom{n}{2}} \right] \approx 1.5083 \binom{n}{3}.$$

Incidentally, the new result of Keevash [4] provides an analog of Lemma 2.6 for general t . Thus we can expect (2.2) as an asymptotic lower bound on 3-laminar families for all integers $n \gg 0$, not just $n = 3^{2^r} + 1$. For $t > 3$, the difficult part is finding good $t-(n, k, 1)$ packings where n is not too much larger than k . This is the fundamental challenge to going further. All we can say in general is that t -laminar families exist with size at least $(1 + \epsilon)\binom{n}{t}$, where $\epsilon > 0$ depends on t .

3. The upper bound: linear programming

We begin with a structural lemma on 2-laminar families.

Lemma 3.1. *Let $\mathcal{F} \subseteq \binom{[n]}{\geq 2}$ be a 2-laminar family and define \mathcal{F}^* to be the set of maximal elements of $\mathcal{F} \setminus \{[n]\}$. Let there be b_i elements of \mathcal{F}^* of cardinality i . Then*

$$|\mathcal{F}| \leq 1 + \sum_{2 \leq k < n} f(k)b_k.$$

PROOF. Every element of $\mathcal{F} \setminus \{[n]\}$ is contained in a unique element of \mathcal{F}^* . For each $K \in \mathcal{F}^*$, the restriction of \mathcal{F} to sets contained in K is a 2-laminar family, say $\mathcal{F}|_K$. But $|\mathcal{F}|_K| \leq f(|K|)$, and so the result follows by summing on K . \square

Next, recall from Proposition 2.1 that the maximal elements of $\mathcal{F} \setminus \{[n]\}$ form a 2-packing on $[n]$. In this way, we can constrain the parameters b_k appearing above.

Lemma 3.2. *Let $([n], \mathcal{B})$ be a 2-packing with b_k blocks of size k , $2 \leq k \leq n$. Then*

$$(3.1) \quad \sum_{2 \leq k \leq n} \binom{k}{2} b_k \leq \binom{n}{2}.$$

Moreover, if $b_m > 0$ and $b_k = 0$ for $m < k \leq n$, we have

$$(3.2) \quad \sum_{2 \leq k \leq m} \binom{k-1}{2} b_k \leq \binom{n-m}{2} + \binom{m-1}{2}.$$

PROOF. The inequality (3.1) follows simply by counting $\binom{[n]}{2}$ in two ways. Next, let $B \in \mathcal{B}$ with $|B| = m$ and count $\binom{[n] \setminus B}{2}$. By the packing condition, every other block of size k intersects $[n] \setminus B$ in at least $k-1$ points. It follows that

$$\sum_{K \in \mathcal{B}, K \neq B} \binom{|K|-1}{2} \leq \binom{n-m}{2},$$

from which (3.2) easily follows. \square

Using just Lemma 3.1 and inequality (3.1), we arrive at a recursive upper bound.

Proposition 3.3. *For all $n > 2$,*

$$\frac{f(n)}{\binom{n}{2}} \leq \frac{1}{\binom{n}{2}} + \max_{2 \leq k < n} \frac{f(k)}{\binom{k}{2}}.$$

PROOF.

$$\frac{f(n)}{\binom{n}{2}} \leq \frac{1}{\binom{n}{2}} + \sum_{2 \leq k < n} \frac{f(k)}{\binom{n}{2}} b_k = \frac{1}{\binom{n}{2}} + \frac{1}{\binom{n}{2}} \sum_{2 \leq k < n} \frac{f(k)}{\binom{k}{2}} \cdot \binom{k}{2} b_k \leq \frac{1}{\binom{n}{2}} + \max_{k < n} \frac{f(k)}{\binom{k}{2}}. \quad \square$$

Corollary 3.4. *For all $n > 2$, $f(n) \leq 2\binom{n}{2}$.*

PROOF. Note that $f(3) = 4$. For $n > 3$, Proposition 3.3 gives

$$\frac{f(n)}{\binom{n}{2}} \leq \frac{f(3)}{\binom{3}{2}} + \sum_{k > 3} \binom{k}{2}^{-1} = 2. \quad \square$$

To use (3.2), we consider m as the maximum cardinality of a set in \mathcal{F}^* . On the one hand, larger values of m contribute a larger $f(m)$ to the bound on $f(n)$. On the other hand, larger values of m constrain the packing via (3.2). To explore this trade-off, we set up a (recursive) linear program.

Define a function $\bar{f} : \mathbb{Z}_{\geq 2} \rightarrow \mathbb{Q}$ by $\bar{f}(2) = 1$, $\bar{f}(3) = 4$ and, for $n > 3$,

$$\bar{f}(n) = 1 + \max_{2 \leq m < n} \text{LP}(n, m),$$

where $\text{LP}(n, m)$ is the solution to the linear program

$$(3.3) \quad \begin{array}{ll} \text{maximize:} & \sum_{2 \leq k \leq m} \bar{f}(k) b_k \\ \text{subject to:} & b_m \geq 1, \quad b_k \geq 0 \text{ for } 3 \leq k < m, \\ & (3.1) \text{ and } (3.2). \end{array}$$

Since this is a fractional relaxation of Lemma 3.1 and the packing constraints hold, we have an upper bound on $f(n)$.

Proposition 3.5. *For all $n > 2$, $f(n) \leq \bar{f}(n)$.*

The linear program (3.3) is somewhat tricky to analyze because of the recursive dependence on \bar{f} . It is probably easiest to just appeal to a long computation which stores \bar{f} in memory.

PROOF OF THEOREM 1.2(b). We compute $\bar{f}(50000) \binom{50000}{2}^{-1} \approx 1.38206$. Then, appealing to Proposition 3.1 we have

$$f(n) \binom{n}{2}^{-1} \leq 1.38206 + \sum_{k > 50000} \binom{k}{2}^{-1} \approx 1.3821. \quad \square$$

It is worth providing some additional details on the computation. First, let us dualize and slightly rewrite (3.3) so that $\text{LP}(n, m)$ is also the solution of the (now two-variable) program

$$(3.4) \quad \begin{aligned} &\text{minimize:} && \binom{n-m}{2}x + \left(\binom{n}{2} - \binom{m}{2}\right)y \\ &\text{subject to:} && x \geq 0, y \geq 1, \text{ and} \\ & && \binom{k-1}{2}x + \binom{k}{2}y \geq \bar{f}(k) \text{ for } k = 3, \dots, m. \end{aligned}$$

A major advantage now is that (3.4) seems to have many redundant constraints. For $k \geq 2$, let η_k denote the halfspace

$$\eta_k = \left\{ (x, y) : \binom{k-1}{2}x + \binom{k}{2}y \geq \bar{f}(k) \right\}.$$

For example, η_2 is defined by $y \geq 1$. As a special definition, let η_1 be the right halfspace defined by $x \geq 0$. As we build a table of $\bar{f}(n)$, we keep track of the intersection

$$\Theta_n := \bigcap_{k=1}^n \eta_k,$$

which is the feasible region for (3.4). When n is incremented, it is a simple check whether $\Theta_n \subseteq \eta_{n+1}$. If so (and this almost always happens), the constraint η_{n+1} is eliminated. For instance, we compute

$$\Theta_{10000} = \eta_1 \cap \eta_2 \cap \eta_3 \cap \eta_7 \cap \eta_{43} \cap \eta_{1807},$$

and we note that these subscripts (starting with 2) are obtained by iterating the map $k \mapsto k^2 - k + 1$. These are the sizes of hypothetical ‘nested’ projective planes. (Owing to the nonexistence of the $2-(43, 7, 1)$ design, though, our construction in Section 2 must shift to use powers of 7. This is essentially what accounts for the small gap between our bounds.) In any case, eliminating constraints at each stage reduces the calculation of $\text{LP}(n, m)$ to a minimum of evaluations of the objective function at the handful of extreme points of Θ_m . This is efficient enough for our purposes to sidestep a more rigorous analysis. However, assuming one can prove the pattern of critical halfspaces continues, a slightly sharper upper bound than Theorem 1.2(b) is possible, namely $f(n) \binom{n}{2}^{-1} \leq 1 + \binom{3}{2}^{-1} + \binom{7}{2}^{-1} + \binom{43}{2}^{-1} + \dots$

The technique in this section can, in principle, extend to $t > 2$, though the bounds likely far exceed what we can construct. A general upper bound of $O(n^t)$ on t -laminar families follows by a similar argument as in Corollary 3.4, and that is enough for an order-of-magnitude determination.

Acknowledgements

Thanks to Richard P. Anstee for helpful discussions and suggestions.

References

- [1] R.P. Anstee, F. Barekat and A. Sali, Small forbidden configurations V: Exact bounds for 4×2 cases, *Studia Sci. Math. Hungarica* 48 (2011), 1–22.
- [2] Y.M. Chee, C.J. Colbourn, A.C.H. Ling, and R.M. Wilson, Covering and packing for pairs. *J. Combin. Theory Ser. A* 120 (2013), 1440–1449.
- [3] J. Cheriyán, T. Jordán and R. Ravi, On 2-coverings and 2-packings of laminar families, in J. Nešetřil (Ed.): *European Symposium 1999*; Lecture Notes in Computer Science 1643 (1999), 510–520.
- [4] P. Keevash, The existence of designs, arXiv preprint <http://arxiv.org/pdf/1401.3665v1.pdf>, 2014.

Peter J. Dukes: Mathematics and Statistics, University of Victoria, Victoria, Canada

E-mail address: `dukes@uvic.ca`